Networks problem sheet 1

Aleksander Kaminski

January 21, 2021

1 Introduction

(1.a): I've already got some programming experience in matlab which guided my decision to never use it again. I will choose to use python, as I don't know it yet.

(1.b): I primarily used website 'learnxiny' [1]to learn the syntax. This was enough for me as a reference to get started. Any other problems I had I consulted the python docs.[2]

(1.d): The format (gml) seems to simply be a text file, firstly defining the nodes and their attributes (label, id) and then a definition of the edges, with edge attributes. The power network didn't have any labels, so I had to manually tell networkx.read_gml to take the id as labels. I used the US Power grid dataset[1], and the Les Misérables[2] dataset. I shall call these power and lesmis, respectively.

(1.e): Firstly, the lesmis dataset, I generated a graph visualisation centered on the node with the highest degree, as shown in Fig. 1. As expected, this node represented the main character of the book. Fig. 2 shows that the second-highest degree nodes had significantly fewer connections than the highest, further showing that the highest-degree character appears with a much more varied audience.

For the power dataset, I firstly used python and networkx to generate a csv file with the degree distribution, then plotted it using R with ggplot2. I also used R to analyze the data. Firstly, I plotted the log-log graphs as shown in Fig. 4. This fit wasn't linear enough to deem correct, and so I moved on to a log- plot as in Fig. 5, which clearly showed a linear behaviour in the subset of degrees [2, 14]. From this, I extracted the exponential behaviour as $\exp(8.25 - 0.528 * degree)$. Fig. 6 shows that there are a large number of "leaf" nodes, namely, nodes which only connect to one other node.

[1] https://learnxinyminutes.com/docs/python/
[2] https://docs.python.org/3/c-api/set.html

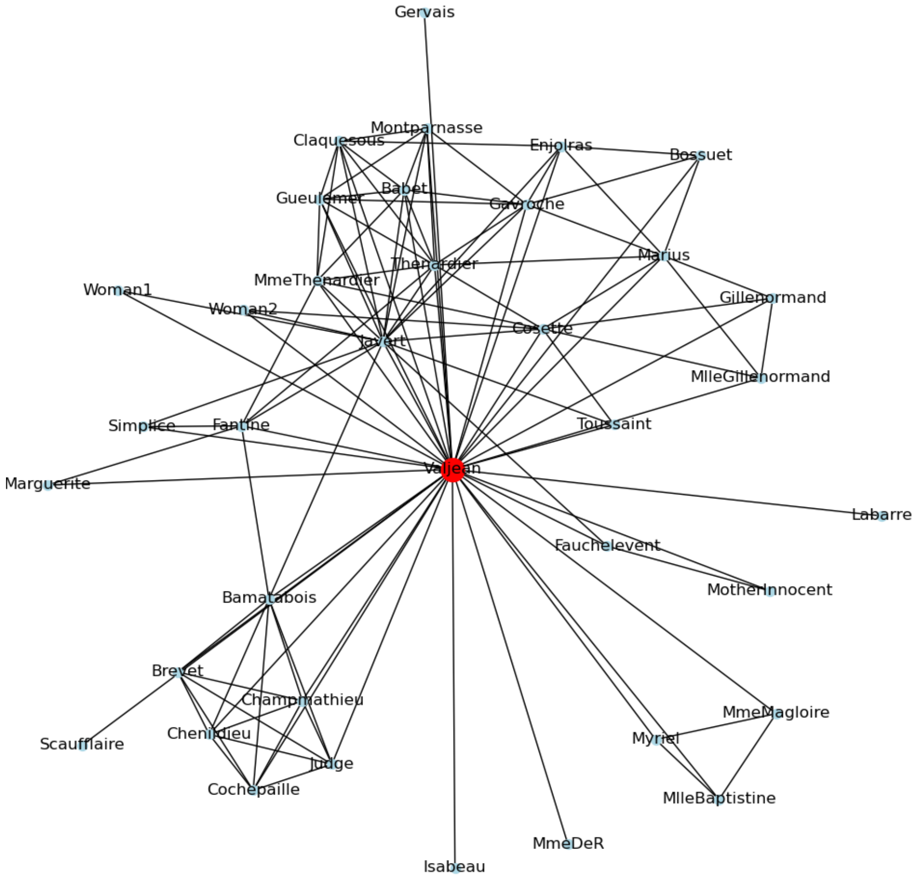

Figure 1: lesmis ego graph

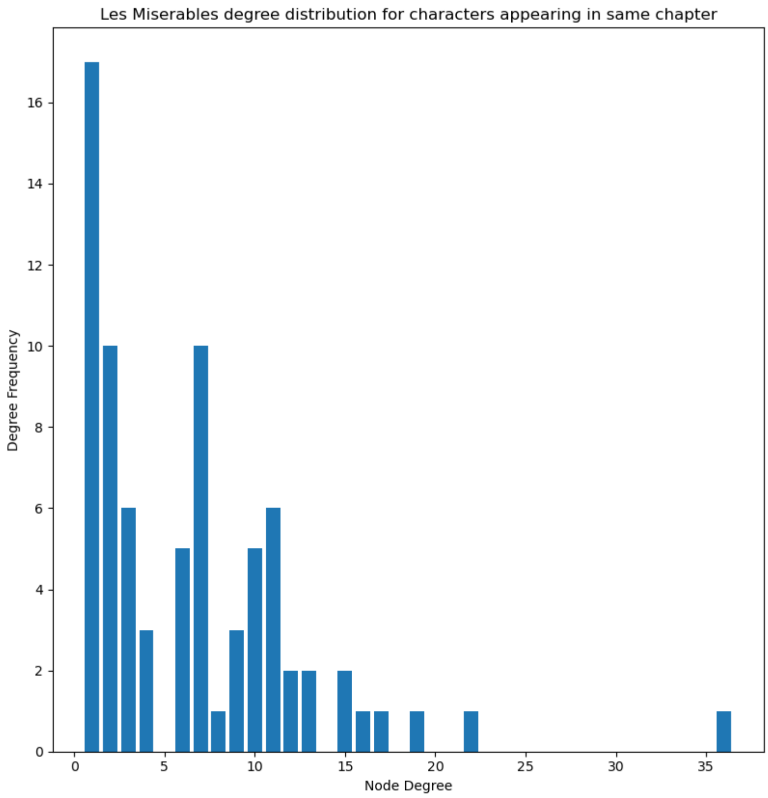

Figure 2: lesmis degree distribution

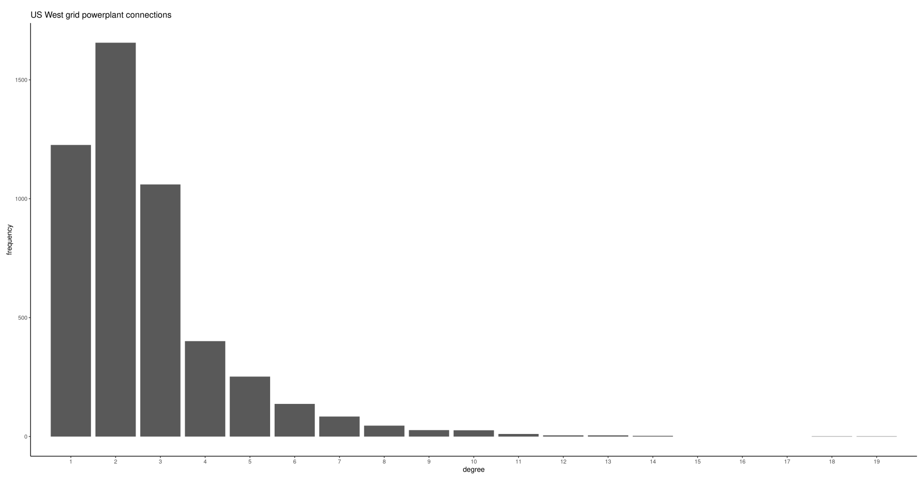

Figure 3: power degree distribution

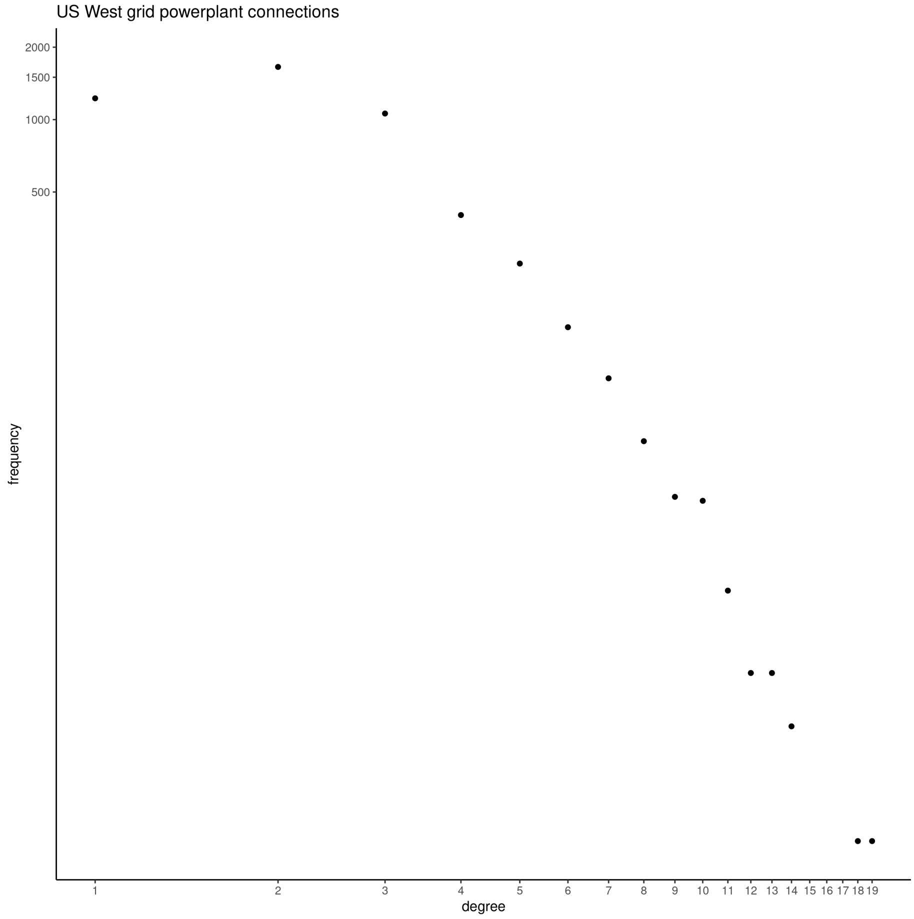

Figure 4: power degree distribution

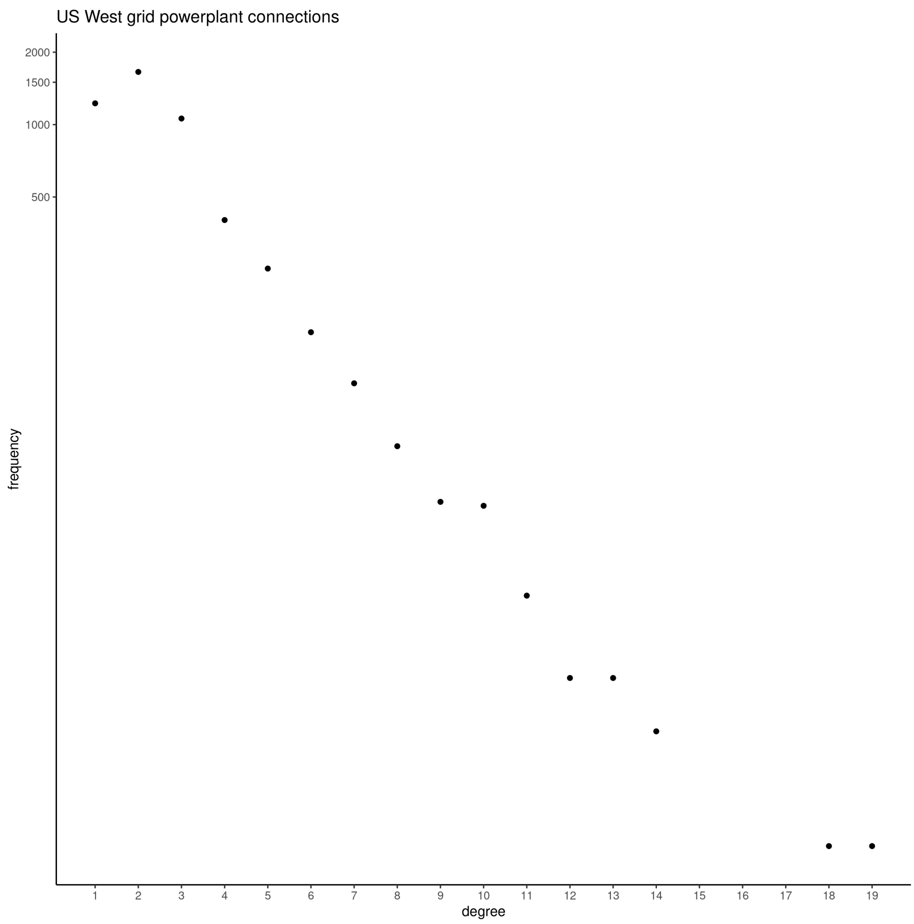

Figure 5: power degree distribution

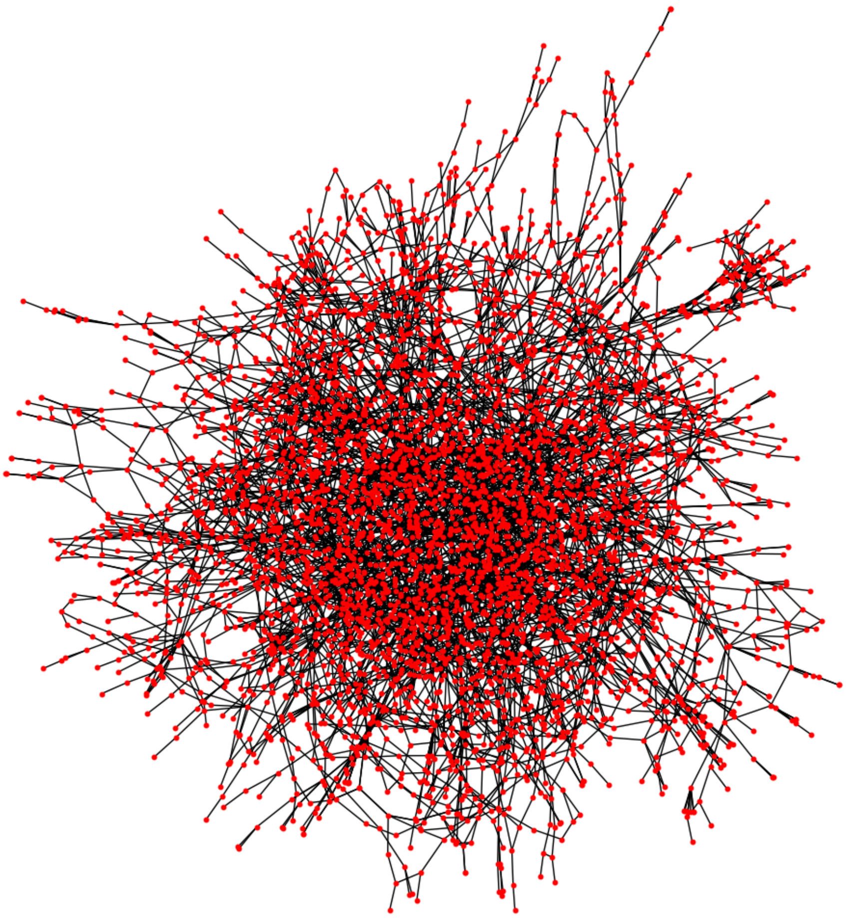

Figure 6: power network visualization

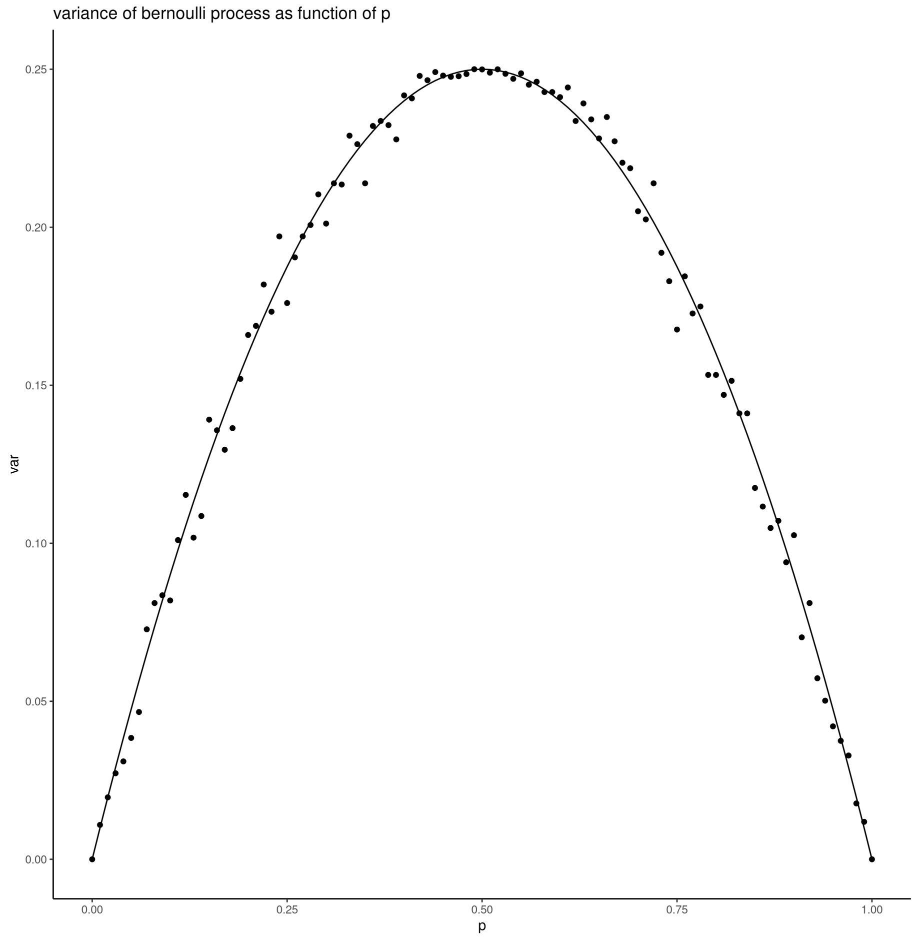

Figure 7: bernoulli process variance, with overlayed $p(1-p)$:

2 Mathematical Toolbox

(2. a) Fig. 7 shows that the computed variance matches the mathematics. The mean was computed to be p.

(2. b) I simulated 10000 random walkers and plotted the resulting position distribution for 1000, 2000 and 3000 samples, overlayed with the gaussian profile

$$p(x;t) = \frac{1}{\sqrt{2\pi Dt}} \exp \frac{-x^2}{4Dt}.$$

With D given by,

$$D = \frac{\langle x^2 \rangle}{2}.$$

The distribution given in the sheet has mean=0

3 Connectedness

References

[1] H. Lietz, *Watts, Duncan J./Strogatz, Steven H. (1998). Collective Dynamics of » Small- World « Networks. Nature 393, S. 440 – 442.*, pp. 551–553. 01 2019.

[2] M. E. J. Newman, "Finding community structure in networks using the eigenvectors of matrices," 2006.